DOCUMENT RESUME

ED 358 144 TM 019 933

AUTHOR Muraki, Eiji; Carlson, James E.
TITLE Full-Information Factor Analysis for Polytomous Item Responses.
SPONS AGENCY National Assessment of Educational Progress, Princeton, NJ.
PUB DATE Apr 93
NOTE 41p.; Paper presented at the Annual Meeting of the American Educational Research Association (Atlanta, GA, April 12-16, 1993).
PUB TYPE Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Computer Graphics; Computer Simulation; Equations (Mathematics); *Estimation (Mathematics); Factor Analysis; Factor Structure; *Item Response Theory; Likert Scales; *Mathematical Models; Multidimensional Scaling; Qualitative Research; *Scoring; *Test Items
IDENTIFIERS EM Algorithm; *Full Information Factor Analysis; *Graded Response Model; Multidimensional Models; Polytomous Items; RESGEN Computer Program

ABSTRACT

A multidimensional item response theory (IRT) model for polytomously scored items, based on the graded response model and using the normal ogive, is developed; and an EM algorithm that may be used to estimate the parameters of the model is also discussed. The model is illustrated through a simulation study in which the polytomous item responses of 6 items and 5,000 respondents were generated by the RESGEN computer program. The relationships between the factor loadings and multidimensional parameters like those of the model of M. D. Reckase (1985) are established. The multidimensional parameters provide useful interpretations of parameters of the multidimensional item response models that can be computed directly from the factor loadings. Estimating the parameters of the multidimensional graded response model is a first step in expanding the application of factor analysis to qualitative data. Four tables and nine computer-generated graphical plots illustrate the analysis. (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☒ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
EIJI MURAKI
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

ED358144

Full-information Factor Analysis

for

Polytomous Item Responses

Eiji Muraki

and

James E. Carlson

Educational Testing Service

This paper was presented at the annual meeting of the American Educational Research Association in Atlanta, April 1993. The work herein was supported under the National Assessment of Educational Progress as administered by the National Center for Education Statistics, Office of Educational Research and Improvement, U. S. Department of Education.

TM019933

The use of Likert items on questionnaires is very popular in sociological and psychological measurement. In the typical analysis of data from such a questionnaire, frequencies of endorsements to points on the scale are tabulated and percentages are computed separately for each item and point of the scale. These descriptive statistics are inconvenient when we try to compare and interpret subjects' responses as measures on the construct we are assessing. Comparisons among items become more cumbersome as the number of items or the number of points on the scale is increased. The points on the scale are frequently assumed to constitute an equal interval scale and subjects' scores are treated as constituting a continuous variable. This assumption is hard to justify if the frequency distribution of scores is highly skewed. Item response theory (IRT) models have provided a solution to problems such as these in other contexts.

Item response theory models applicable to scales made up of dichotomously-scored items measuring one proficiency dimension have been developed and are now in widespread use. The distinct feature of IRT models as compared to other statistical models for categorical variables, such as logistic and log-linear models, is the inclusion of latent traits as variables. The model that is the subject of this paper is based on two extensions of the basic IRT model. Models that can incorporate polytomously-scored items have been proposed and used by several researchers (Andrich, 1978, 1982, 1988; Bock, 1972; Masters, 1982; Muraki, 1990, 1992a; Samejima, 1969, 1972). Bock and Aitkin (1981) extended IRT models for dichotomously-scored items to the multidimensional case (several proficiency dimensions) and developed an EM algorithm (Dempster, Laird, & Rubin, 1977) to estimate the parameters of the model based on the normal ogive. McKinley & Reckase (1983) proposed a logistic-distribution-based multidimensional model. In this paper a multidimensional IRT model for polytomously-scored items, based on Samejima's graded response model and using the normal ogive, is developed, and an EM algorithm that may be used to estimate the parameters of the model, is also discussed.

Development of the Model

Bock, Gibbons, and Muraki (1988) assume that the interaction of person i and item j results in a response process variable, y_{ij}, that is a linear combination of M latent traits. Using vector notation in which $\underline{\theta}$ is an M-dimensional vector of latent traits (common factors) and $\underline{\alpha}$ a vector of factor loadings:

$$\underline{\theta}'_i = (\theta_{1i}, \theta_{2i}, \ldots, \theta_{Mi})$$
$$\underline{\alpha}'_j = (\alpha_{j1}, \alpha_{j2}, \ldots, \alpha_{jN})$$

we may write this combination as

$$y_{ij} = \underline{\alpha}'_j \underline{\theta}_i + e_{ij}$$
$$= \alpha_{j1}\theta_{1i} + \alpha_{j2}\theta_{2i} + \ldots + \alpha_{jM}\theta_{Mi} + e_{ij} \qquad [1]$$

where ϵ_{ij} is an unobserved random variable (a unique factor in factor analytic terms) which is assumed to be distributed N(0, σ_j^2). Conventionally in factor analysis it is assumed that the distribution of $\underline{\theta}$ is N($\underline{0}$, **I**) and that y_j is distributed with mean 0 and variance 1. Hence the unique variance is

$$\sigma_j^2 = 1 - \sum_{m=1}^{M} \alpha_{jm}^2 = 1 - \underline{\alpha}'\underline{\alpha} \qquad [2]$$

Classical factor analysis for continuous variables is based on the assumption that the response process is directly observable. In contrast, the factor analytic model for categorical variables is based on the assumption that the response process variable, y_{ij}, is latent and realized into a vector of polytomous item responses for J items,

$$\underline{W}_i = (W_{i1}, W_{i2}, \ldots, W_{iJ})$$

according to the psychological mechanism

$$W_{ij} = k \quad if \quad \gamma_{j,k-1} \leq y_{ij} < \gamma_k$$
$$\gamma_{j0} = -\infty$$
$$\gamma_{jK} = \infty$$
$$(k = 1, 2, \ldots, K)$$

where γ_{jk} is a threshold parameter associated with category k of a K-category Likert-type item, j. The process generates a categorical response of k for person i to item *j* when y_{ij} equals or exceeds the threshold, $\gamma_{j,k-1}$, but does not exceed the threshold γ_{jk}. Assuming a normal ogive model the probability of categorical response k by person i to item *j* given his/her M-dimensional latent trait is expressed as

$$P(W_{ij} = k \mid \underline{\theta}_i) = \int_{\gamma_{j,k-1}}^{\gamma_{jk}} \exp\left[-\frac{1}{2}\left(\frac{y_{ij} - \underline{\alpha}'_j \underline{\theta}_i}{\sigma_j} \right)^2 \right] dy \qquad [3]$$

Introducing a change of variable, we may re-write the item response model in **[3]** by defining

$$t_{ij} = \frac{y_{ij} - \boldsymbol{\alpha}_j'\boldsymbol{\theta}_i}{\sigma_j} \qquad [4]$$

Then

$$dy_{ij} = \sigma_j dt_{ij} \qquad [5]$$

and when

$$y_{ij} = \gamma_{jk} \qquad [6]$$

then

$$t_{jk} = \frac{\gamma_{jk} - \boldsymbol{\alpha}_j'\boldsymbol{\theta}_i}{\sigma_j} \qquad [7]$$

and when

$$y_{ij} = \gamma_{j,k-1} \qquad [8]$$

then

$$t_{j,k-1} = \frac{\gamma_{j,k-1} - \boldsymbol{\alpha}_j'\boldsymbol{\theta}_i}{\sigma_j} \qquad [9]$$

Following Bock, Gibbons, and Muraki (1988) we define slope and item-category parameters:

$$a_{jm} = \frac{\alpha_{jm}}{\sigma_j} \qquad [10]$$

$$b_{jk} = -\frac{\gamma_{jk}}{\sigma_j} \qquad [11]$$

$$(k = 1, 2, \ldots, K\text{-}1)$$

and define the following functions:

$$Z_{jk}(\underline{\theta}) = \underline{a}'_j\underline{\theta} + b_{jk} \qquad (k = 1, 2, \ldots, K-1) \qquad [12]$$

$$P^*_{jk}(\underline{\theta}) = \sum_{c=1}^{k} P_{jc}(\underline{\theta}) \qquad [13]$$

$$P^*_{j0} = 0. \qquad [14]$$

$$P^*_{jK} = 1. \qquad [15]$$

$$\phi(t) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) \qquad [16]$$

then the model in [3] may be written as

$$P_{jk}(\underline{\theta}) = P^*_{jk}(\underline{\theta}) - P^*_{j,k-1}(\underline{\theta}) \qquad [17]$$

which can be written as

$$P_{jk}(\underline{\theta}) = \int_{-\infty}^{z_{jk}(\underline{\theta})} \phi(t)\,dt - \int_{-\infty}^{z_{j,k-1}(\underline{\theta})} \phi(t)\,dt \qquad (k = 1, 2, \ldots, K) \qquad [18]$$

For k=1 the last term is zero and for k=K the first term is the definite integral of a probability density function from negative infinity to infinity and therefore equals one. Hence we may write these two cases as

$$P_{j1}(\underline{\theta}) = P^*_{j1}(\underline{\theta})$$
$$P_{jK}(\underline{\theta}) = 1.0 - P^*_{j,K-1}(\underline{\theta}) \qquad [19]$$

Equation 17 is referred to as the operating characteristic of the graded response model by Samejima (1972).

Once we obtain the estimates of slope and item-category parameters, we can compute the corresponding factor loadings and threshold parameters from relationships derived from [2], [10], and [11], as follows:

$$\alpha_{jm} = \frac{a_{jm}}{\xi_j} \qquad [20]$$

and

$$\gamma_{jk} = -\frac{b_{jk}}{\xi_j} \qquad [21]$$

where

$$\xi_j^2 = 1+\sum_{m=1}^{M} a_{jm}^2 = 1+\underline{a}'\underline{a} \qquad [22]$$

Interpretation of Parameters in Multidimensional IRT Models

Reckase (1985, 1986) developed "multidimensional" parameters that serve as an aid in understanding the nature of items generated according to the multidimensional logistic model and Carlson (1987) further elaborated on their meaning. Here we apply similar ideas to items generated by multidimensional models for polytomously-scored items and, in particular, the normal ogive model in [17] and [18].

Following Reckase's (1985, 1986) definition, the multidimensional discrimination of a polytomously-scored item can be defined as

$$\eta_j^2 = \sum_{m=1}^{M} a_{jm}^2 = \underline{a}_j'\underline{a}_j \qquad [23]$$

and the multidimensional item-category parameter can be defined as

$$\beta_{jk} = -\frac{b_{jk}}{\eta_j} \qquad [24]$$

Reckase (1985) found the direction of steepest slope of the response surface can expressed in terms of angles with the latent trait axes as

$$\cos\omega_{jm} = \frac{a_{jm}}{\eta_j} = \lambda_{jm} \qquad [25]$$

The slope in the direction specified by angles ω_{jm} is at its maximum when the item response surface (IRS) of $P_{jk}^{+}(\theta)$ crosses the .5 probability hyperplane.

The item-category parameter of the cumulative probability in [13] is the proficiency value at the point of .5 probability of the kth or below kth categorical response where

$$\sum_{m=1}^{M} a_{jm}\theta_m = -b_{jk} \qquad [26]$$

If we denote a particular dimension as m' ($1 \leq m' \leq M$), we obtain the following equation because of the symmetric relation:

$$\theta_m = \frac{\theta_{m'}}{a_{jm'}} a_{jm} \qquad [27]$$

Equations [26] and [27] can be solved from the points in the proficiency dimensions locating the item-category parameter:

$$\begin{aligned} \theta_{m'} &= -a_{jm'}\frac{b_{jk}}{\eta_j^2} \\ &= \frac{a_{jm'}}{\eta_j}\beta_{jk} \\ &= \beta_{jk}\cos\omega_{jm'} \end{aligned} \qquad [28]$$

From [25] and [28], we obtain the following equations:

$$a_{jm} = \eta_j \cos \omega_{jm} \qquad [29]$$

and

$$\tau_{jkm} = \beta_{jk} \cos \omega_{jm} \qquad [30]$$

Since the direction cosine is $\cos(\omega_{jm})$ for both the multidimensional discrimination η_j and the multidimensional difficulty β_{jk}, they must reside along the same axis. The slope parameter a_{jm} is the mth coordinate of the point where the slope of IRS is the steepest. The value τ_{jkm} indicates the mth coordinate of the point on the line. In other words, it is the projection of β_{jk} onto the mth θ dimension. Thus, the parameter, τ_{jkm}, can be referred to the mth component of the multidimensional item-category parameter, β_{jk}.

The concepts discussed above can be best explained in the context of two dimensions. Figures 1 and 2 show a two-dimensional response surface and contour plot, respectively, for $P_{jk}^{+}(\theta)$ in [13] with *a*-parameters of .8 and .6 and b_{jk}-parameter of .5. Note that the locus of all points for which $\theta_1=0.$ defines a line in the response surface (Figure 1) which is a unidimensional item characteristic curve above the θ_2 axis. Similarly, for $\theta_2=0.$ there is a two-parameter unidimensional item characteristic curve over the θ_1 axis. The η_j for the example item j in the figures is therefore equal to 1.0 and the β_{jk} is -0.5. The direction cosines of the line of maximum slope in reference to θ_1 and θ_2 are .8 and .6, respectively. Since these θ axes are orthogonal, $\cos^2\omega_{j1} + \cos^2\omega_{j2}=1$.

****
Insert Figures 1 and 2 about here
****

Consider the geometry in Figure 3. The θ_1 and θ_2 are two orthogonal axes representing the two dimensions underlying item j. The linear combination in **[12]** indicates the combination of the two θs that the item can be considered to be measuring:

$$\underline{a}_j'\underline{\theta} = a_{j1}\theta_1 + a_{j2}\theta_2$$

The equation of the line defining the direction of measurement, as is shown in Figure 3, can be written as

$$\theta_2 = \frac{a_{j2}}{a_{j1}}\,\theta_1$$

The two a parameters are the lengths of two sides (a_{j1} = OC and a_{j2} = OA) of a right triangle whose hypotenuse (OB) is equal to Reckase's multidimensional discrimination parameter, η_j. The angle BOC is ω_{j1} and the angle AOB is its orthogonal complement, ω_{j2}. The two τ parameters are the lengths of two sides (τ_{j11} = OF and τ_{j12} = OD) of a right angle whose hypotenuse (OE) is equal to the multidimensional difficulty parameter (β_{jk}). The multidimensional item response surface of the cumulative probability is a two-dimensional surface above the θ_1-θ_2 plane (Figure 1), and a slice through that surface along the line of measurement of the item is a unidimensional item characteristic curve for an item with discrimination parameter equal to η_j and item-category parameter equal to β_{jk} (Carlson, 1987).

****
Insert Figure 3 about here
****

By using the multidimensional parameters defined above, the model in [12] can be rewritten as

$$\begin{aligned} Z_{jk}(\underline{\theta}) &= a_{j1}\theta_1 + a_{j2}\theta_2 + \ldots + a_{jM}\theta_M + b_{jk} \\ &= \eta_j(\cos\omega_{j1}\theta_1 + \cos\omega_{j2}\theta_2 + \ldots + \cos\omega_{jM}\theta_M - \beta_{jk}) \\ &= \eta_j(\lambda_{j1}\theta_1 + \lambda_{j2}\theta_2 + \ldots + \lambda_{jM}\theta_M - \beta_{jk}) \\ &= \eta_j(\theta_j^* - \beta_{jk}) \end{aligned} \qquad [31]$$

The unidimensional latent trait θ_j^* is a composite of M dimensional latent traits θ_m (m=1, 2, ..., M). Reckase pointed out that an item, although being scaled in a multidimensional context, can be considered to be measuring along a single dimension. That dimension is a linear combination of the uncorrelated theta dimensions. The test containing the item may, however, be multidimensional if it consists of items that measure along different directions in the theta space. Since it is assumed that the distribution of $\underline{\theta}$ is $N(\underline{0},\underline{I})$ and the sum of squared direction cosines in the orthogonal space is 1., the variance of θ_j^* is also 1.

The model can be also expressed as a linear combination of factor loadings and thresholds:

$$\begin{aligned} Z_{jk}(\underline{\theta}) &= \xi_j(\alpha_{j1}\theta_1 + \alpha_{j2}\theta_2 + \ldots + \alpha_{jM}\theta_M - \gamma_{jk}) \\ &= \frac{\theta_j^{**} - \gamma_{jk}}{\sigma_j} \end{aligned} \qquad [32]$$

Unlike the composite latent trait θ_j^* in [31], the variance of a linear combination of the weighted individual θ_m, θ_j^{**}, in [32] is not 1. The variance of this composite of latent traits is called the communality, h_j^2. Since

$$Z_{jk}(\underline{\theta}) \sim N(-\frac{\beta_{jk}}{\eta_j}, \eta_j^2) \qquad [33]$$

$$Z_{jk}(\underline{\theta}) \sim N(-\frac{\gamma_{jk}}{\sigma_j}, \frac{h_j^2}{\sigma_j^2}) \qquad [34]$$

and

$$\xi_j^2 = 1+\eta_j^2, \qquad [35]$$

the communality is expressed by the parameters in [31] as

$$h_j^2 = \frac{\eta_j^2}{1+\eta_j^2} \qquad [36]$$

The multidimensional discrimination and direction cosines are expressed by the parameters in the model [32] as

$$\eta_j = \frac{h_j}{\sqrt{1-h_j^2}} = \frac{h_j}{\sigma_j} \qquad [37]$$

and

$$\lambda_{jm} = \frac{a_{jm}}{\eta_j} = \frac{\alpha_{jm}}{h_j} \qquad [38]$$

As seen in [10] and [37], the multidimensional discrimination can be interpreted in the same way as the conversion of factor loadings to the slope parameters. Furthermore, the direction cosine can be also computed as the ratio of the factor loadings to the square root of the communality.

To summarize, the multidimensional parameters defined for the McKinley-Reckase M2PL model (McKinley & Reckase, 1983; Reckase, 1985) can be adopted for the multidimensional graded response model in [18] because the cumulative probability of each categorical response of kth or below kth, $P_{jk}^{+}(\underline{\theta})$, is essentially a dichotomous item response model and the logistic function is only an approximate form of the normal ogive function. The difference is that the multidimensional polytomous item response model yields a set of K-1 IRSs of $P_{jk}^{+}(\underline{\theta})$ rather than one such surface for the dichotomous item response model. These IRSs are parallel along the line defined by a set of direction cosines, λ_{jm}, m=1,2, ..., M. The probability of the specific middle category k, $P_{jk}(\underline{\theta})$ is defined by subtracting $P_{j,k-1}^{+}(\underline{\theta})$ from $P_{jk}^{+}(\underline{\theta})$. Since these IRSs of cumulative probabilities are parallel, the multidimensional parameters are still meaningful even for specific middle categories.

Bock, Gibbons, and Muraki (1988) established the relationships between the parameters of the factor analysis model and the parameters of the item response model. In this paper, we have established the relationships between the factor loadings and multidimensional parameters like those in Reckase's model. The multidimensional parameters can provide useful interpretations of parameters of the multidimensional item response models. These parameters can be computed directly from the factor loadings.

Figures 4 through 7 show the IRSs of polytomously-scored three-category item 1 with parameters: $a_{11}=1.0$, $a_{12}=1.5$, $b_{11}=-0.8$, and $b_{12}=1.2$. In Figure 4, the IRS of $P_{11}(\underline{\theta})$ is plotted. This is the same probability as $P_{11}^{+}(\underline{\theta})$, as given in [19]. In Figure 5, the IRS of the second cumulative probability $P_{12}^{+}(\underline{\theta})$ is plotted. These two IRSs are parallel to each other. The model probability of the second categorical response, $P_{12}(\underline{\theta})= P_{12}^{+}(\underline{\theta})-P_{11}^{+}(\underline{\theta})$, is computed and plotted in Figure 6. Finally, the model probability of the third categorical response, $P_{13}(\underline{\theta})= 1.0 - P_{12}^{+}(\underline{\theta}b)$, is plotted in Figure 7. If the width of the item-category parameters is shortened, the probability of the middle categorical response is uniformly decreased. Consequently, the IRS becomes flatter. In Figure 8, the IRS of the middle categorical response of item 2 with the same parameter values as item 1 except that $b_{21}=0.8$ is shown. Since the width of the item-category parameters becomes narrower given the same slope parameters, the IRS is pushed to downward. We can expect the observed response frequencies of the second category to be smaller for item 2 than for item 1. The shape of the IRS does not change if we rotate the original IRS. In Figure 9, the IRS of the second categorical response, $P_{32}(\underline{\theta})$, with slope parameters $a_{31}=1.5$ and $a_{32}=1.0$ is plotted.

****

Insert Figures 4, 5, 6, and 7 here

****

Parameter Estimation

Let U_{ijk} represent an element in the matrix of the observed response pattern i. $U_{ijk}=1$ if item j is rated by the ith respondent in the kth category of a Likert scale, otherwise $U_{ijk}=0$. By the principle of local independence (Birnbaum, 1968) the conditional probability of a response pattern i, given $\underline{\theta}$, for K response categories and J items, as denoted by a response matrix (U_{jk}), is the joint probability:

$$P_i[(U_{jk}) \mid \underline{\theta}] = \prod_{j=1}^{J} \prod_{k=1}^{K} [P_{jk}(\underline{\theta})]^{U_{ijk}} \qquad [39]$$

For examinees randomly sampled from a population with a multivariate normal distribution of the latent trait variable, $\phi(\underline{\theta})$, the marginal probability of the observed response pattern i is

$$P_i[(U_{jk})] = \int_{\underline{\theta}} P_i[(U_{jk}) \mid \underline{\theta}] \phi(\underline{\theta}) d\underline{\theta} \qquad [40]$$

If an examinee responds to J items with K categories, his/her response pattern i can then be assigned to one of K^J mutually exclusive patterns. Let r_i represent the number of examinees observed in such a pattern i, and let N be the total number of examinees sampled from the population. Then r_i is multinominally distributed with parameters N and $P_i(U_{jk})$. This probability can be interpreted as the likelihood function of the parameters a_{jm} and b_{jk}:

$$L = \frac{N!}{\prod_{i=1}^{K^J} r_i!} \prod_{i=1}^{K^J} [P_i[(U_{jk})]]^{r_i} \qquad [41]$$

Taking the natural logarithm of Equation 41 yields

$$\ln L = \ln N! - \sum_{i=1}^{K^J} \ln r_i! + \sum_{i=1}^{K^J} r_i \ln P_i[(U_{jk})] \qquad [42]$$

Bock and Aitkin (1981) applied the EM algorithm (Dempster et al., 1977) to estimate the parameters for each item individually, and then repeated the iteration process over J items until the estimates of all the items became stable to the required number of decimal places. This is in contrast to the Fisher-scoring procedure of Bock and Lieberman (1970). The qth cycle of the iterative process can be expressed as

$$\underline{\nu}^{q} = \underline{\nu}^{q-1} + V^{-1}\underline{t} \qquad [43]$$

for each item j. The vector of estimates $\underline{\nu}$ represents the model parameters. The orders of parameter vector $\underline{\nu}$ and gradient vector $\underline{t}$ are both M+K-3, and the order of information matrix V is (M+K-3) x (M+K-3). The information matrix is the negative expectation of the matrix of second derivatives. When the number of response categories is K, only K-1 item category parameters can be specified. In addition, two extreme ends of parameter values need to be fixed to estimate slopes. Furthermore, because the covariance between any categories that differ by more than two points is simply 0, the partitioned information matrix for the category parameter estimation becomes a tridiagonal symmetric matrix (Muraki, 1990).

The likelihood equations for $\hat{a}_{jm}$ and $\hat{b}_{jk}$ can be derived from the first derivatives of Equation 42 with respect to each parameter, and respectively set to 0.

With respect to a_{jm}, the likelihood in [42] can be differentiated as

$$\frac{\partial \ln L}{\partial a_{jm}} = \sum_{i=1}^{K^J} \frac{r_i}{P_i[(U_{jk})]} \int_{\underline{\theta}} P_i[(U_{jk}) \mid \underline{\theta}] \sum_{k=1}^{K} \frac{\partial [P_{jk}(\underline{\theta})]^{v_{ijk}}}{\partial a_{jm}} \frac{\phi(\underline{\theta})\, d\underline{\theta}}{[P_{jk}(\underline{\theta})]^{v_{ijk}}} \qquad [44]$$

Now let the observed score patterns be indexed by l=1,2,..., S, where $S \leq \min(N, K^J)$. If the number of examinees with response pattern l is denoted by r_l, then

$$\sum_{l=1}^{S} r_l = N \qquad [45]$$

The first derivative of the likelihood function in [44] can be approximated by using the Gauss-Hermite quadrature, such that

$$\frac{\partial \ln L}{\partial a_{jm}} \approx \sum_{l=1}^{S} \sum_{f_M=1}^{F_M} \cdots \sum_{f_2=1}^{F_2} \sum_{f_1=1}^{F_1} \frac{r_l L_l(X) A(X_{f_1}) A(X_{f_2}) \cdots A(X_{f_M})}{\tilde{P}_l} \sum_{k=1}^{K} \frac{\partial [P_{jk}(X)]^{v_{ijk}}}{\partial a_{jm}} \frac{1}{[P_{jk}(X)]^{v_{ijk}}} \qquad [46]$$

where

$$\tilde{P}_l = \sum_{f_M=1}^{F_M} \cdots \sum_{f_2=1}^{F_2} \sum_{f_1=1}^{F_1} \prod_{j=1}^{J} \prod_{k=1}^{K} [P_{jk}(X)]^{U_{ljk}} A(X_{f_1}) A(X_{f_2}) \cdots A(X_{f_M}) \qquad [47]$$

and

$$L_l(X) = \prod_{j=1}^{J} \prod_{k=1}^{K} [P_{jk}(X)]^{U_{ljk}} \qquad [48]$$

In Equation 46, $A(X_f)$ is the weight of the Gauss-Hermite quadrature, and X_f is the quadrature point (Stroud & Secrest, 1966). The quadrature weight $A(X_f)$ is approximately the standard normal probability density at the point X_f for each dimension. Because U_{ljk} can take only two possible values, 1 and 0, the element of the gradient vector $\underline{t}_a$ can be written as

$$t_{a_{jm}} = \sum_{f_M=1}^{F_M} \cdots \sum_{f_2=1}^{F_2} \sum_{f_1=1}^{F_1} \sum_{k=1}^{K} \frac{\bar{r}_{jkf_1f_2\cdots f_M}}{P_{jk}(X)} \frac{\partial P_{jk}(X)}{\partial a_{jm}} \qquad [49]$$

where

$$\bar{r}_{jkf_1f_2\cdots f_M} = \sum_{l=1}^{S} \frac{r_l L_l(X) A(X_{f_1}) A(X_{f_2}) \cdots A(X_{f_M}) U_{ljk}}{\tilde{P}_l} \qquad [50]$$

and $\bar{r}_{jkf_1f_2\cdots f_M}$ is called the provisional expected frequency of the kth categorical response of item j at the fth quadrature point.

The item category parameter, b_{jk}, is contained in both $P_{jk}(\underline{\theta})$ and $P_{j,k+1}(\underline{\theta})$ as shown in [18]. The first derivative of Equation 39 with respect to b_{jk} is given by

$$\frac{\partial P_i[(U_{jk})|\underline{\theta}]}{\partial b_{jk}} = P_i[(U_{jk})|\underline{\theta}] \left[\frac{U_{ijk}}{P_{jk}(\underline{\theta})} - \frac{U_{i,j,k+1}}{P_{j,k+1}(\underline{\theta})}\right] \frac{\partial P^*_{jk}(\underline{\theta})}{\partial b_{jk}} \qquad [51]$$

Therefore, the element of the gradient vector $\underline{t}_b$ is numerically computed as

$$t_{b_{jk}} = \sum_{f_M=1}^{F_M} \cdots \sum_{f_2=1}^{F_2} \sum_{f_1=1}^{F_1} \left[\frac{\bar{r}_{jkf_1f_2\cdots f_M}}{P_{jk}(X)} - \frac{\bar{r}_{j,k+1,f_1f_2\cdots f_M}}{P_{j,k+1}(X)}\right] \frac{\partial P^*_{jk}(X)}{\partial b_{jk}} \qquad [52]$$

The elements of the information matrix are given by

$$V_{a_{jm}a_{jm'}} = \sum_{t_M=1}^{F_M} \cdots \sum_{t_2=1}^{F_2} \sum_{t_1=1}^{F_1} \bar{N}_{jt_1t_2\cdots t_M} \sum_{k=1}^{K} \frac{1}{P_{jk}(X)} \frac{\partial P_{jk}(X)}{\partial a_{jm}} \frac{\partial P_{jk}(X)}{\partial a_{jm'}} \qquad [53]$$

$$V_{b_{jk}b_{jk}} = \sum_{t_M=1}^{F_M} \cdots \sum_{t_2=1}^{F_2} \sum_{t_1=1}^{F_1} \bar{N}_{jt_1t_2\cdots t_M} \left[\frac{1}{P_{jk}(X)} + \frac{1}{P_{j,k+1}(X)} \right] \left[\frac{\partial P^*_{jk}(X)}{\partial b_{jk}} \right]^2 \qquad [54]$$

$$V_{b_{jk}b_{j,k+1}} = -\sum_{t_M=1}^{F_M} \cdots \sum_{t_2=1}^{F_2} \sum_{t_1=1}^{F_1} \bar{N}_{jt_1t_2\cdots t_M} \frac{1}{P_{j,k+1}(X)} \frac{\partial P^*_{jk}(X)}{\partial b_{jk}} \frac{\partial P^*_{j,k+1}(X)}{\partial b_{j,k+1}} \qquad [55]$$

and, when $|k-k'| \geq 2$,

$$V_{b_{jk}b_{jk'}} = 0 \qquad [56]$$

and finally

$$V_{a_{jm}b_{jk}} = \sum_{t_M=1}^{F_M} \cdots \sum_{t_2=1}^{F_2} \sum_{t_1=1}^{F_1} \bar{N}_{jt_1t_2\cdots t_M} \left[\frac{1}{P_{jk}(X)} \frac{\partial P_{jk}(X)}{\partial a_{jm}} - \frac{1}{P_{j,k+1}(X)} \frac{\partial P_{j,k+1}(X)}{\partial a_{jm}} \right] \frac{\partial P^*_{jk}(X)}{\partial b_{jk}} \qquad [57]$$

The algorithm presented above was implemented in POLYFACT (Muraki, 1993). The POLYFACT is a hybrid computer program of PARSCALE (Muraki & Bock, 1993) and TESTFACT (Wilson, Wood, & Gibbons, 1984). The program computes the factor loadings by the principal factor analysis based on the product moment correlation matrix, treating item responses as a continuous variable. Because the factors of the principal factor analysis are orthogonal, their loadings are suitable for the full-information solution after conversion to slopes. Slope estimates based on the full-information method are then converted again into factor loadings. The resulting full-information factor loadings are then rotated orthogonally to the varimax criterion (Kaiser, 1958) and, with the varimax solution as target, rotated obliquely by the promax method (Hendrickson & White, 1964).

Simulation Study

The polytomous item responses of six items and 5000 respondents were generated by the RESGEN (Muraki, 1992b) computer program. All six items have three categorical

responses. The values of the original parameters are presented in Table 1. The values of $a=1.2$ and $a=0.0$ correspond to $\alpha=0.768$ and $\alpha=0.0$, respectively. The communality is the same for all six items ($h_j^2=0.590$, $j=1,2,\ldots,6$).

****
Insert Table 1 here
****

The three-dimensional solutions were obtained. Five quadrature points for each dimension were used for numerical integration. Therefore, 5^3 total points were used for integration. The precision for convergence was set 0.001. Nineteen EM cycles were needed to reach convergence. The estimated slope parameters were converted to factor loadings, and the communality for each item was computed. The factor loadings were then rotated with a varimax criteria and they are presented in Table 2. The slope parameters were recovered from the varimax factor loading, and they are shown in Table 3. The item-category parameters are also presented in Table 3. Reckase's multidimensional discrimination parameter (MDP) and direction cosines were computed based on the varimax factor loadings and are shown in Table 4.

****
Insert Tables 2, 3, and 4 here
****

All of the slope parameters were underestimated, and consequently the varimax factor loadings are lower than the original ones. The item-category parameters are also underestimated. The underestimation of these parameters may be eased by increasing the number of quadrature points. We are currently investigating this possibility. Nevertheless, the factor structure of the original simulated data is recovered in the estimated parameters. The process of conversion from the slope parameters into factor loadings and the rotation of the factor loadings is an efficient way to study the results of the analysis. We can then recover the slope parameters and compute the multidimensional discrimination parameters and the direction cosines.

For further research, we are planning the analysis of real data sets of polytomous item responses. The test of chi-square fit is essential to investigation of the appropriateness of the multidimensional item response models. The interpretation of the multidimensional parameters in terms of classical factor analysis is also needed. Since increasing the dimensionality requires an exponential increment in computational time, we need to investigate adjusting the number of quadrature points to economize the estimation process. In this research, we attempted to estimate the parameters of the multidimensional graded

response model. We succeeded in this basic task. This is the first step in expanding the application of the factor analysis to qualitative data.

REFERENCES

Andrich, D. (1978). A rating formulation for ordered response categories. Psychometrika, 43, 561-573.

Andrich, D. (1982). An extension of the Rasch model for ratings providing both location and dispersion parameters. Psychometrika, 47, 105-113.

Andrich, D. (1988). A general form of Rasch's extended logistic model for partial credit scoring. Applied Measurement in Education, 1, 363-378.

Birnbaum, A. (1968). Some latent trait models and their use in inferring an examinee's ability. In F. M. Lord, & M. R. Novick, Statistical theories of mental test scores (pp. 397-479). Reading, MA: Addison Wesley.

Bock, R. D. (1972). Estimating item parameters and latent ability when responses are scored in two or more nominal categories. Psychometrika, 37, 29-51.

Bock, R. D., & Aitkin, M. (1981). Marginal maximum likelihood estimation of item parameters: Application of an EM algorithm. Psychometrika, 46, 443-459.

Bock, R. D., Gibbons, R., & Muraki, E. (1988). Full-information item factor analysis. Applied Psychological Measurement, 12, 261-280.

Bock, R. D., & Lieberman, M. (1970). Fitting a response model for n dichotomously scored items. Psychometrika, 35, 179-197.

Carlson, J. E. (1987). Multidimensional item response theory estimation: A computer program. (Research Report 87-19) Iowa City, IA: The American College Testing Program.

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. Journal of the Royal Statistical Society (Series B), 39, 1-38.

Hendrickson, A. E., & White, P. O. (1964). PROMAX: A quick method for rotation to oblique simple structure. British Journal of Mathematical and Statistical Psychology, 17, 65-70.

Kaiser, H. F. (1958). The varimax criterion for analytic rotation in factor analysis. Psychometrika, 23, 187-200.

Masters, G. N. (1982). A Rasch model for partial credit scoring. Psychometrika, 47, 149-174.

McKinley, R. L., & Reckase, M. D. (1983). An extension of the two-parameter logistic model to the multidimensional latent space. (Research Report ONR 83-2). Iowa City, IA: The American College Testing Program.

Muraki, E. (1990). Fitting a polytomous item response model to Likert-type data. Applied Psychological Measurement, 14, 59-71.

Muraki, E. (1992a). A generalized partial credit model: Application of an EM algorithm. Applied Psychological Measurement, 16, 159-176.

Muraki, E. (1992b). RESGEN: Item Response Generator. Princeton, NJ: Educational Testing Service.

Muraki, E. (1993). POLYFACT (Unpublished computer Program). Princeton, NJ: Educational Testing Service.

Muraki, E., & Bock, R. D. (1991). PARSCALE: Parameter scaling of rating data. Chicago, Ill.: Scientific Software, Inc.

Reckase, M. D. (1985). The difficulty of test items that measure more than one ability. Applied Psychological Measurement, 9, 401-412.

Reckase, M. D. (1986, April). The discriminating power of items that measure more than one dimension. Paper presented at the meeting of the American Educational Research Association, San Francisco.

Samejima, F. (1969). Estimation of latent ability using a response pattern of graded scores. Psychometrika Monograph Supplement, 34 (Monograph Number 17).

Samejima, F. (1972). A general model for free-response data. Psychometrika Monograph Supplement, 37 (Monograph Number 18).

Wilson, D., Wood, R., & Gibbons, R. D. (1984). TESTFACT: Test scoring, item statistics, and item factor analysis. Chicago: Ill.: Scientific Software, Inc.

Table 1

Original Model Parameters

	Slope Parameter			Item-Category Parameter	
Item	a_1	a_2	a_3	b_1	b_2
1	1.2	0.0	0.0	-1.8	0.8
2	1.2	0.0	0.0	-0.8	1.8
3	0.0	1.2	0.0	-1.8	0.8
4	0.0	1.2	0.0	-0.8	1.8
5	0.0	0.0	1.2	-1.8	0.8
6	0.0	0.0	1.2	-0.8	1.8

Table 2

Varimax Factor Loadings and Communality

	Factor Loadings			Communality	
Item	α_1	α_2	α_3	h^2	h
1	0.709	0.002	-0.007	0.503	0.709
2	0.705	-0.002	-0.014	0.497	0.705
3	0.006	0.699	0.021	0.489	0.699
4	-0.006	0.698	0.003	0.487	0.698
5	0.012	-0.017	-0.723	0.523	0.723
6	0.010	-0.009	-0.729	0.532	0.729

Table 3

Estimated Model Parameters

	Slope Parameter			Item-Category Parameter	
Item	a_1	a_2	a_3	b_1	b_2
1	1.005	0.003	-0.010	-1.204	0.492
2	0.994	-0.003	-0.020	-0.546	1.211
3	0.008	0.978	0.029	-1.121	0.506
4	-0.008	0.975	0.004	-0.522	1.161
5	0.017	-0.025	-1.047	-1.149	0.549
6	0.015	-0.013	-1.065	-0.507	1.191

Table 4

Reckase's Multidimensional Parameters

Item	MDP η	Direction Cosines λ_1	λ_2	λ_3
1	1.005	1.000	0.003	-0.010
2	0.994	1.000	-0.003	-0.020
3	0.978	0.009	1.000	0.030
4	0.975	-0.009	1.000	0.004
5	1.047	0.017	-0.024	-1.000
6	1.065	0.014	-0.012	-1.000

Figure 1. Item Response Surface for the Multidimensional Graded Response Model with Parameters $a_1 = 0.8$, $a_2 = 0.6$, and $b_1 = 0.5$.

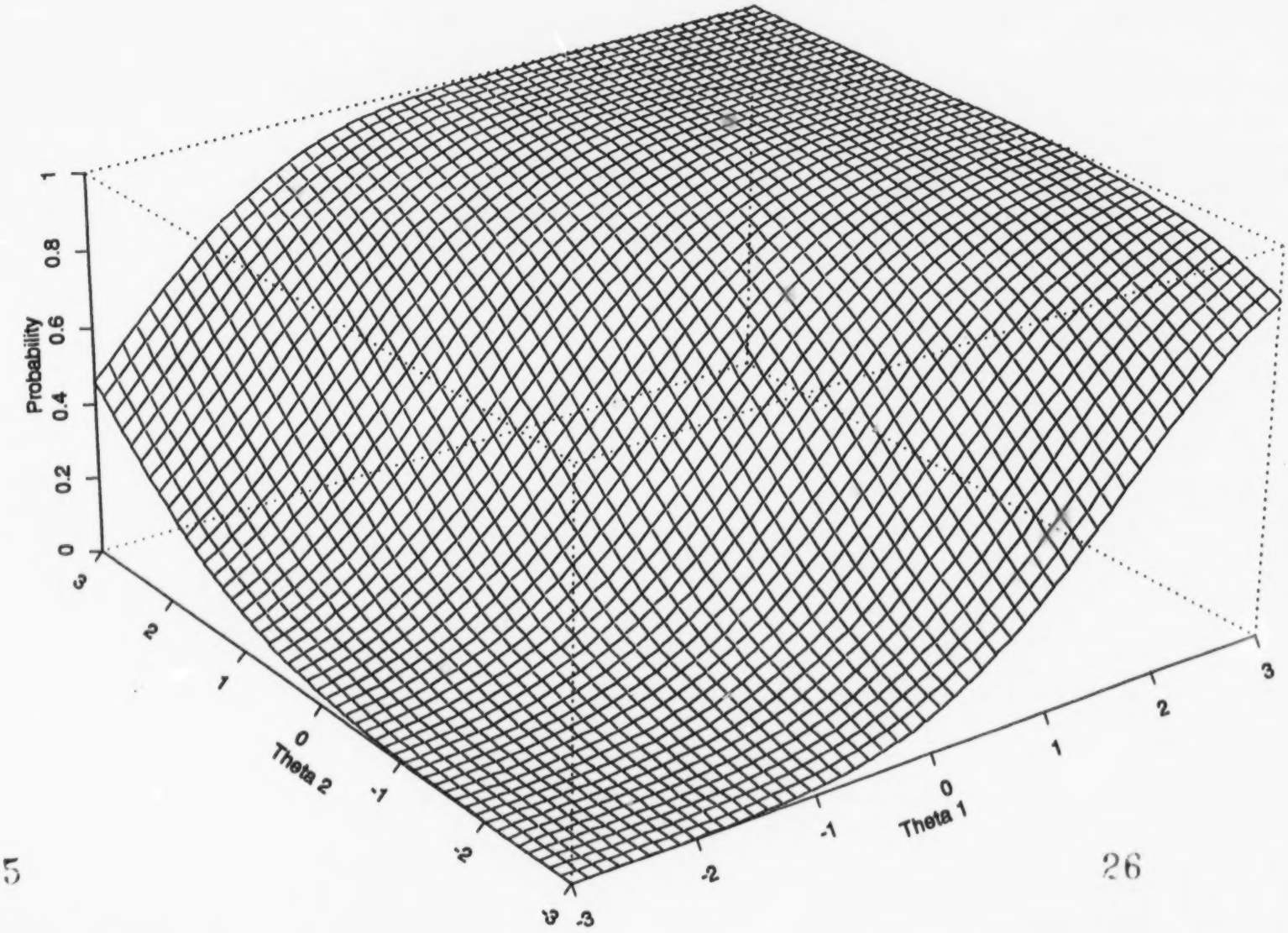

Figure 2. Contour Plot for the Multidimensional Graded Response Model with Parameters $a_1 = 0.8$, $a_2 = 0.6$, and $b_1 = 0.5$.

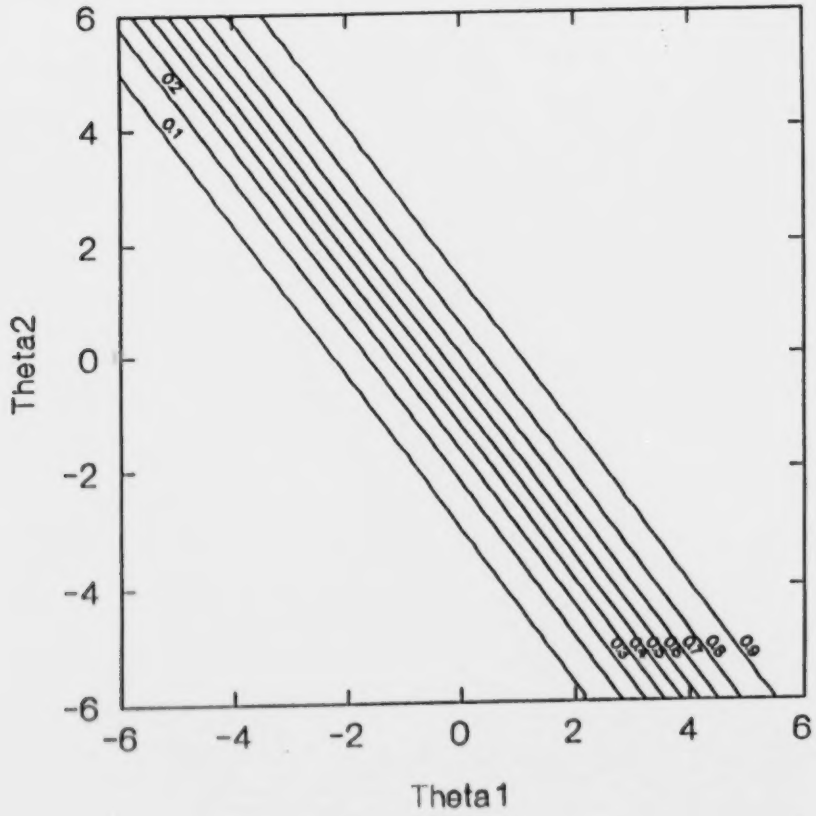

Figure 3. Multidimensional Discrimination for a Two-Dimensional Space.

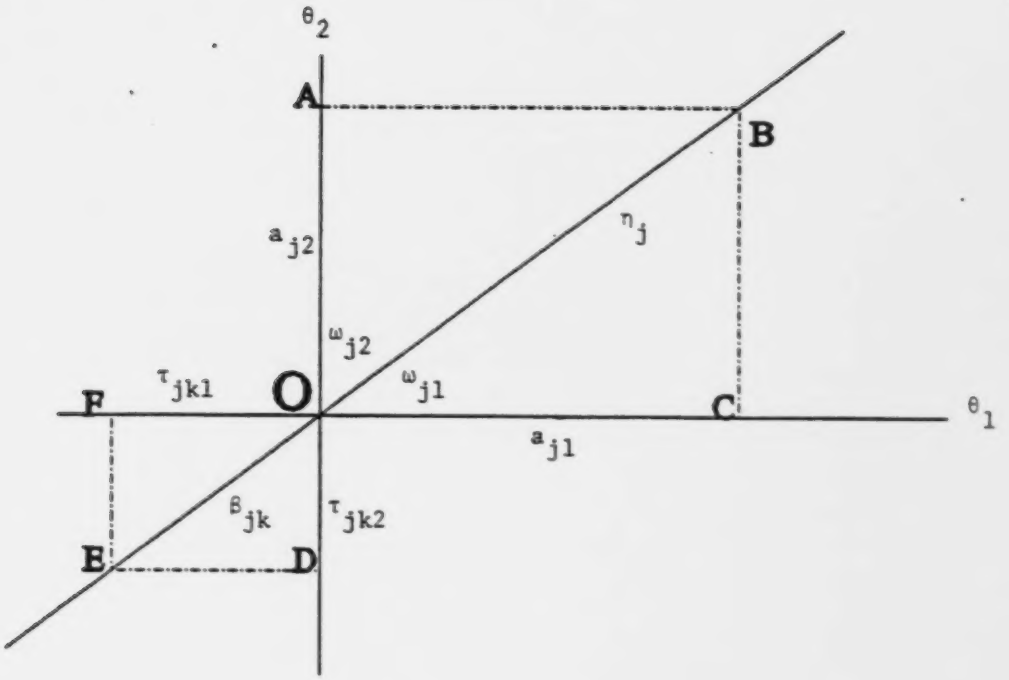

Figure 4. Item Response Surface for the Multidimensional Graded Response Model $P_{11}(\underline{\theta})$ with Parameters $a_{11} = 1.0$, $a_{12} = 1.5$, $b_{11} = -0.8$, and $b_{12} = 1.2$.

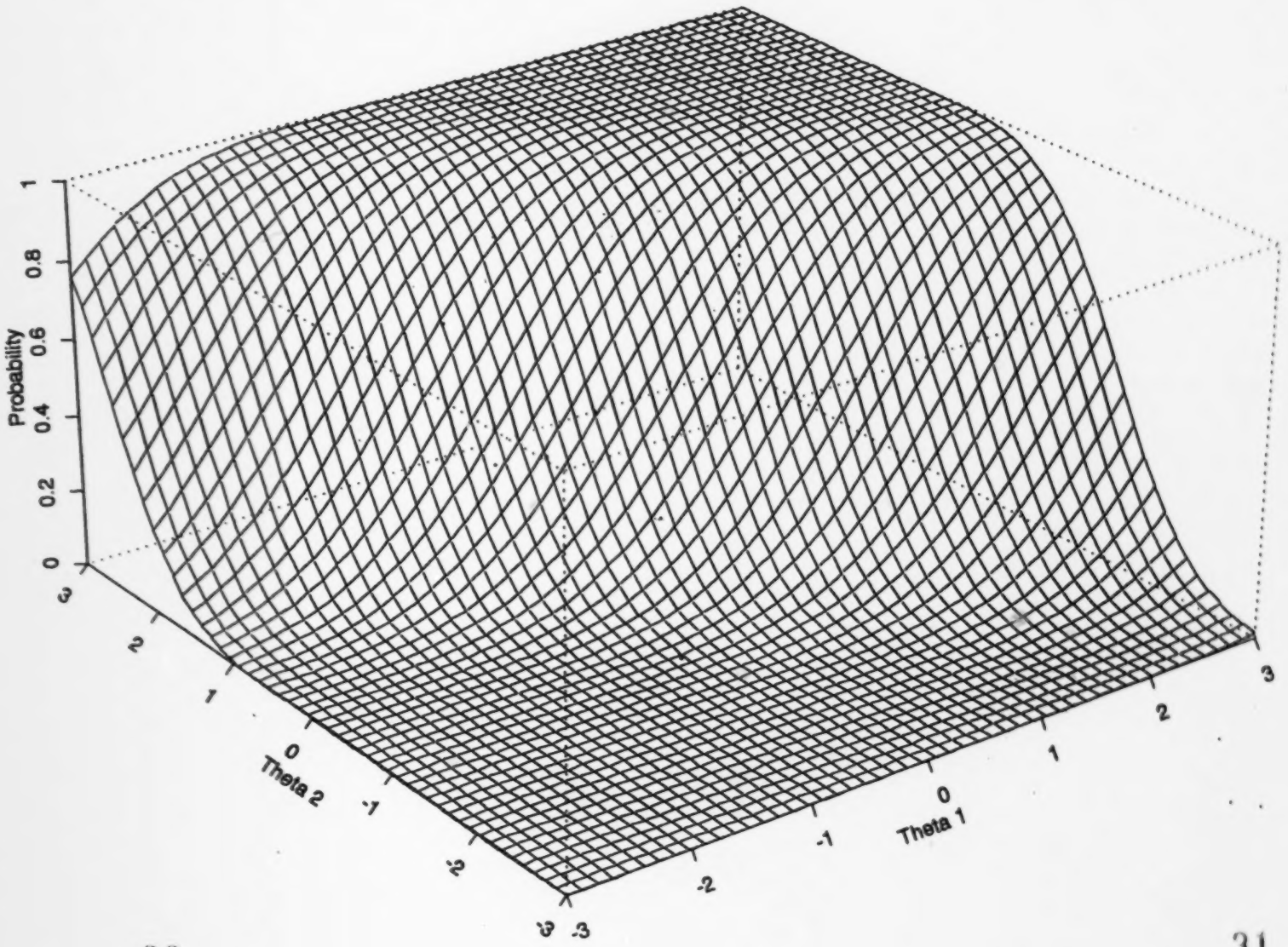

Figure 5. Item Response Surface for the Multidimensional Graded Response Model $P_{12}^{*}(\underline{\theta})$ with Parameters $a_1 = 1.0$, $a_2 = 1.5$, $b_{11} = -0.8$, and $b_{12} = 1.2$.

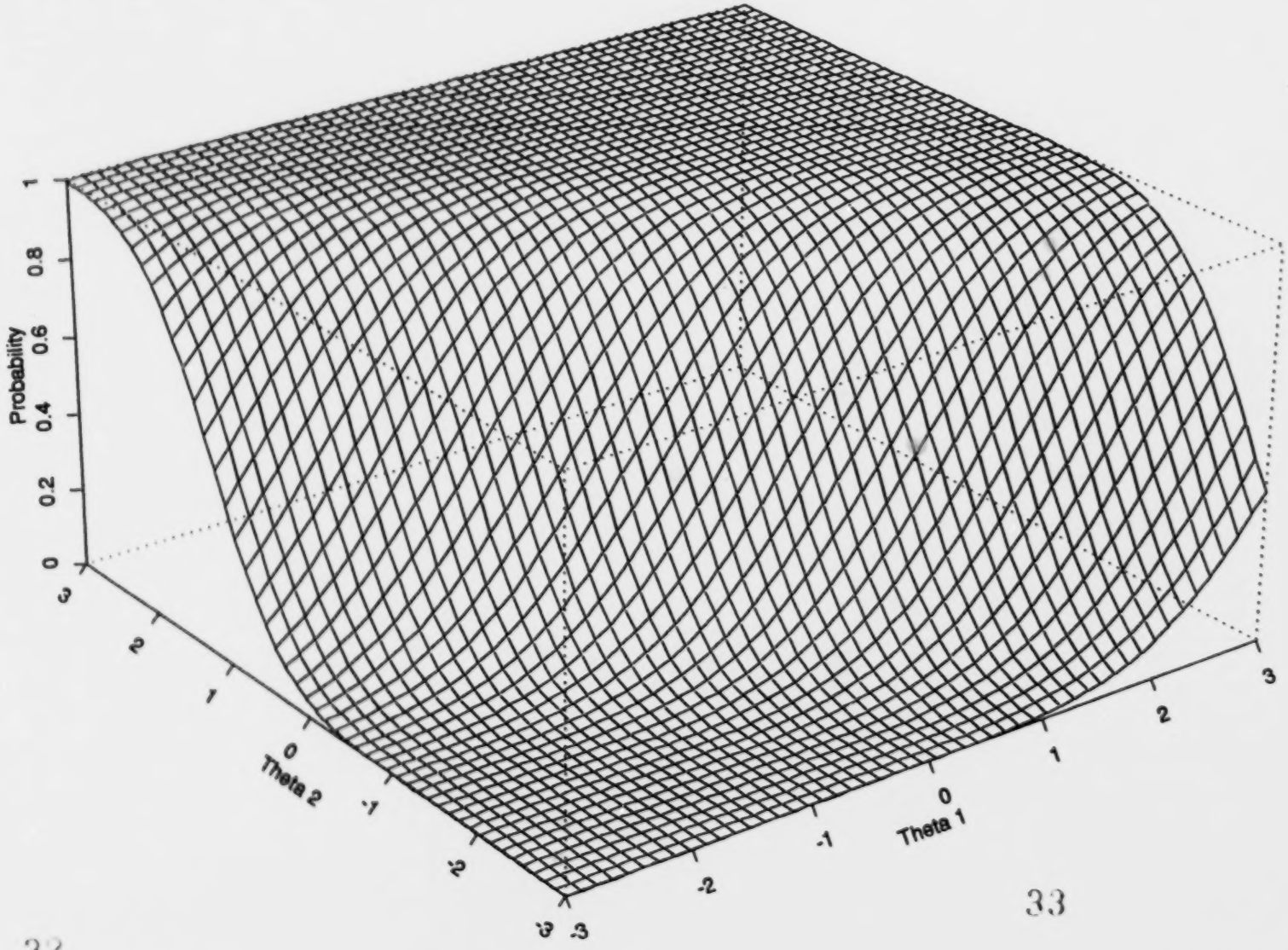

Figure 6. Item Response Surface for the Multidimensional Graded Response Model $P_{12}(\theta)$ with Parameters $a_1 = 1.0$, $a_2 = 1.5$, $b_{11} = -0.8$, and $b_{12} = 1.2$.

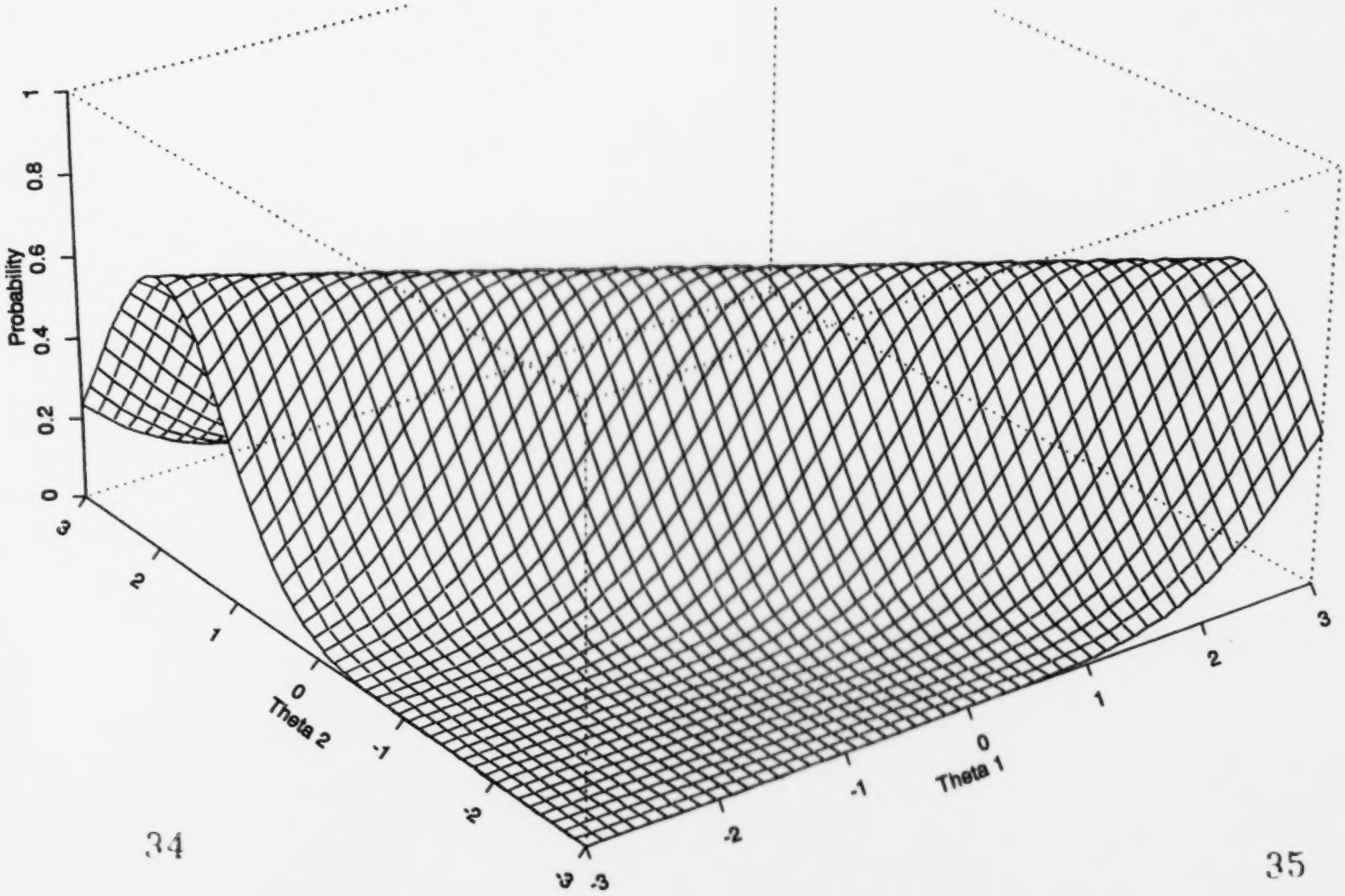

Figure 7. Item Response Surface for the Multidimensional Graded Response Model $P_{13}(\underline{\theta})$ with Parameters $a_1 = 1.0$, $a_2 = 1.5$, $b_{11} = -0.8$, and $b_{12} = 1.2$.

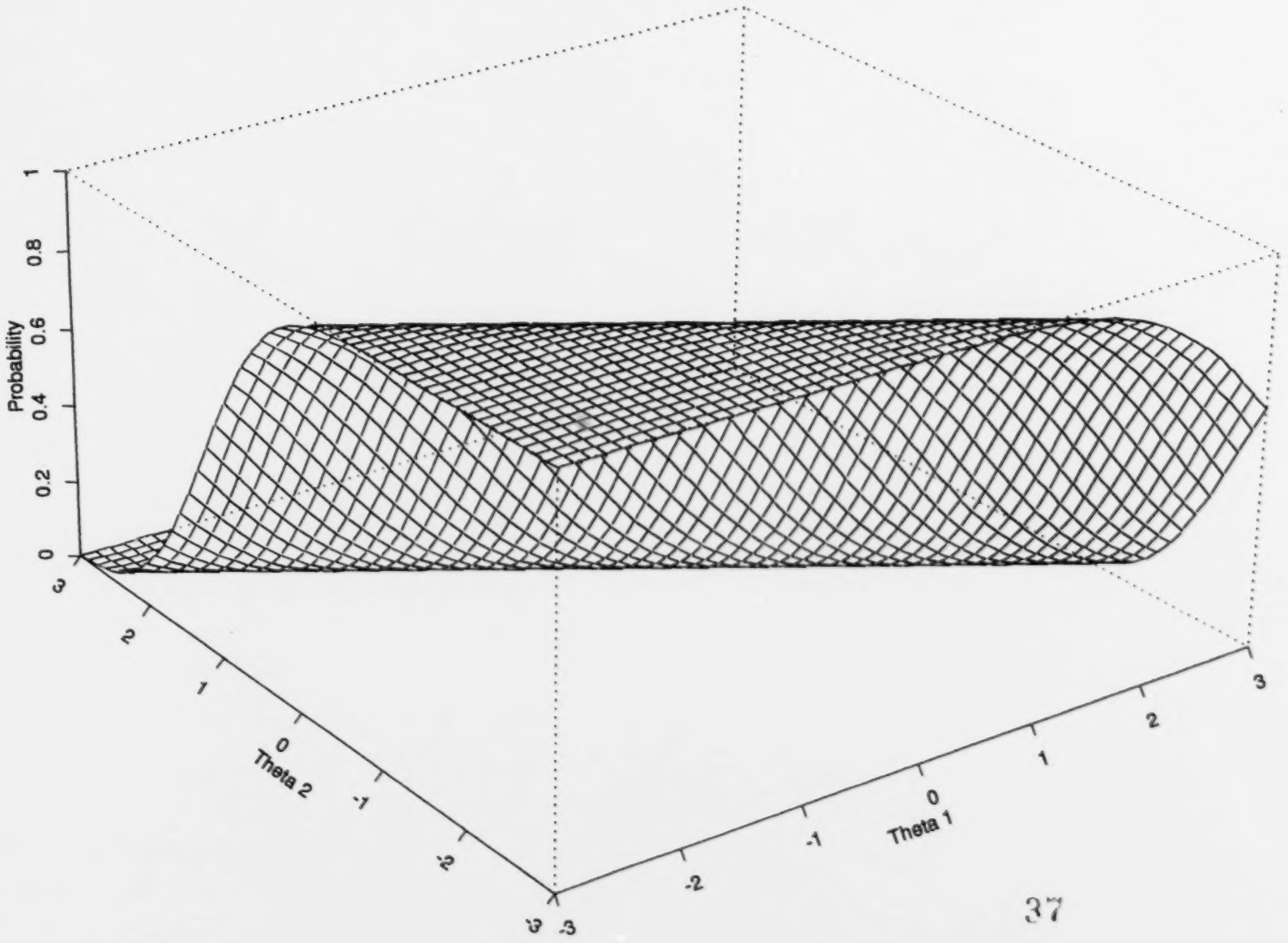

Figure 8. Item Response Surface for the Multidimensional Graded Response Model $P_{22}(\theta)$ with Parameters $a_1 = 1.0$, $a_2 = 1.5$, $b_{21} = -0.8$, and $b_{22} = 1.2$.

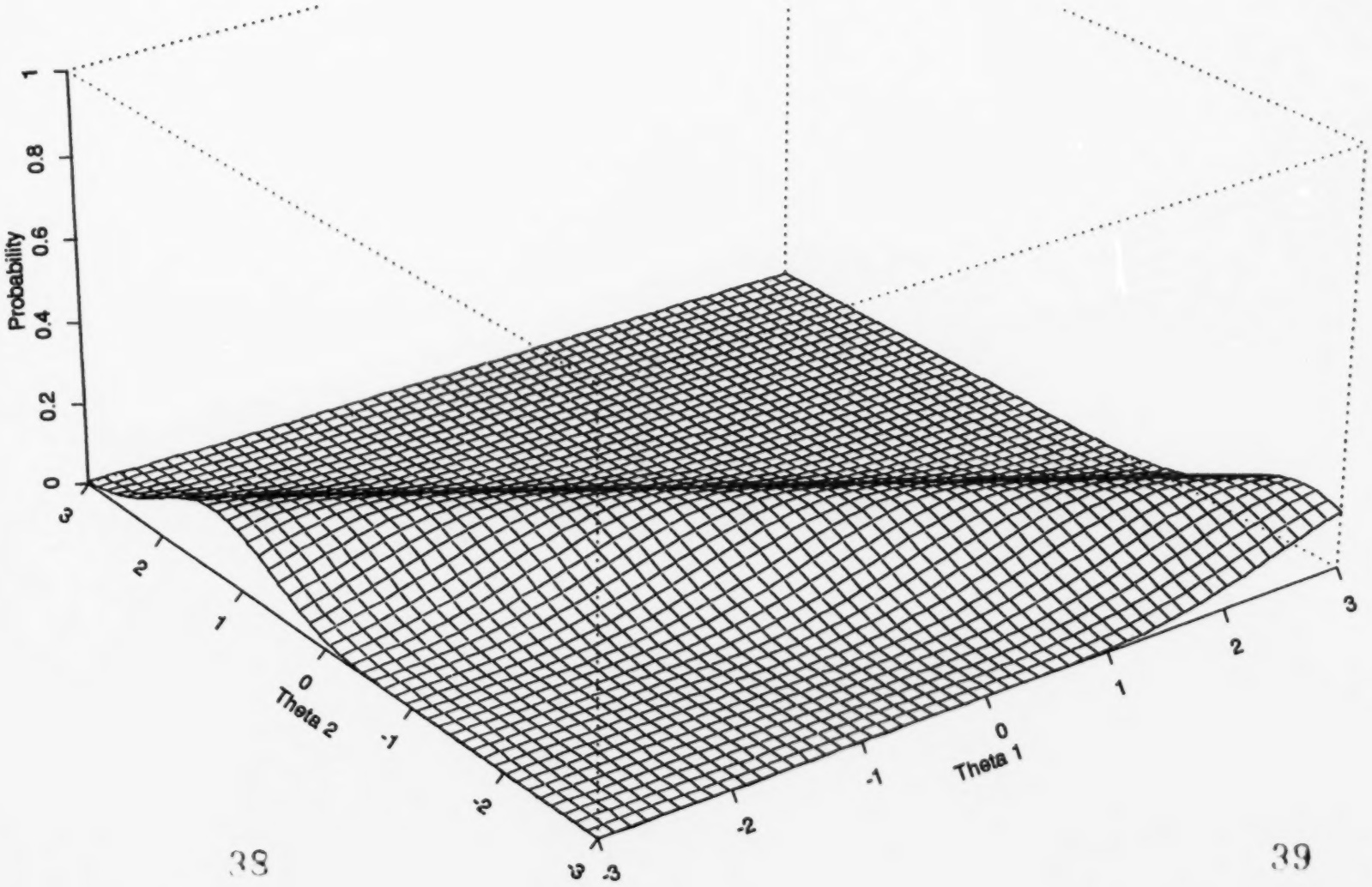

Figure 9. Item Response Surface for the Multidimensional Graded Response Model $P_{32}(\theta)$ with Parameters $a_1 = 1.0$, $a_2 = 1.5$, $b_{31} = -0.8$, and $b_{32} = 1.2$.